[redacted]

D/FBIS
DD/FBIS
1 C/E&PS
C/LRB
CMO
C/AG
2 C/PROD
C/OPS
C/ADMIN
201 FILE
3 EXEC. REG.

ILLEGIB

10 December 1981

NOTE FOR: Executive Officer, DDS&T

Attached for your information is a copy of our response to [redacted] STAT query as to whether FBIS (or anyone) should sponsor a seminar on Machine Translation. As you can see, we decline but suggest another approach (para. 7).

[redacted] STAT

Director

Attachment
Memo to [redacted] IR&DC/IC Staff STAT
on Seminar on Machine Translation

[redacted] 10Dec81 STAT

Distribution:
Orig. - Addressee, w/att
1 - D/FBIS Chrono, wo/att
1 - [redacted] Correspondence, wo/att STAT
1 - FBIS Registry, wo/att

Org + Mgt Translation Seminar

9 DEC 1981

D/FBIS
DD/FBIS
1 C/E&PS B
C/LRS
CMC
C/AG
C/PROD
C/CPS
C/ADMIN
201 FILE
2 EXEC. REG.

MEMORANDUM FOR: [redacted] STAT
Executive Secretary, IR&DC/IC Staff

FROM: [redacted] STAT
Director, Foreign Broadcast Information Service

SUBJECT: Seminar on Machine Translation

1. Your recent visit to FBIS has had the salutary effect of focusing our minds again on the current state of the art in machine translation and stimulating a review of our understanding of that subject. We remain in intermittent contact with those who are active in it--the Chief of our Joint Publications Research Service is our designated representative at meetings, both government and nongovernment, which touch upon it--and so we have felt reasonably certain that our lines are open. Nevertheless we found your visit a useful occasion to take stock again of where we stand.

2. FBIS/Production Group's involvement in Machine Translation (through its predecessor, the Foreign Documents Division) dates back to the decade of the late fifties/early sixties, when it was the primary actor in CIA's investment of [redacted] STAT to develop a workable machine translation system. The CIA investment was but a small portion of the nearly [redacted] STAT invested by the U.S. Government as a whole. After nearly a decade of repeated promises but little delivery, the U.S. Government sponsors requested the National Academy of Sciences/National Research Council to review the status and promise of machine translation and make appropriate recommendations. The results of that review, published under the title "LANGUAGE AND MACHINES: Computers in Translation and Linguistics" led to the abandonment of CIA sponsorship of machine translation.

3. For the next decade, FBIS monitored the progress of machine translation and provided limited financial support to the Chinese-English Translation Association (CETA), a less ambitious project with greater benefits. By 1975, the Director, FBIS, felt a need for a more systematic review of the state of the art of machine translation, and at his urging in 1976 FBIS' Production Group sponsored a 2-day seminar in Rosslyn which was attended by some 100 participants and observers, including 15 guest speakers. To set up the seminar, a

SUBJECT: Seminar on Machine Translation

contract in the amount of [redacted] was given to [redacted] President of MRM Inc., 811 Connecticut Avenue, Kensington, MD, who also heads CETA. [redacted] was known by FBIS to have good contacts in both the machine translation and machine-assisted translation fields, and he put together a group of speakers and participants who were considered to be representative of the business, academic, and government communities. A copy of [redacted] bid is attached for your information as Tab A. STAT STAT STAT

4. One of the benefits of this exercise was the summation of the seminar's results by 10 of the key participants (copy attached as Tab B). Very briefly summarized, their conclusions were: a) that machine translation had limited application to volume translation activities, and then only to activities of very constrained scope utilizing very limited vocabularies; b) that successful application of MT to activities of a more generalized nature--political, literary, or artistic--was not possible without a significant breakthrough in the field of artificial intelligence (in effect, creation of a thinking machine); c) therefore, organizations seeking to modernize their translation activities should concentrate on the field of machine-assisted translations; i.e., hardware and software aids to translation. Taking its cue from these conclusions, FBIS sponsored another seminar in 1978; this one geared to Machine Assisted Translation (MAT), but including some of the same players as the first one--people who still hoped to convince the U.S. Government that their machine translation systems would solve the linguist's problems. The second seminar confirmed the validity of our earlier conclusions: there had been no significant breakthrough in the state of the art in MT, and MAT was still the most likely avenue of success.

5. FBIS has since concentrated on machine aids to translation. It has a test-bed system now nearing operational use for the compilation, updating, and dissemination of foreign language glossaries for the use of its [redacted] language officers at Headquarters, its [redacted] independent contract translators scattered throughout the United States, and its field network of some [redacted] foreign national radio monitors. It also has plans, now well advanced, for computerization of its Consolidated Translations Survey, a data bank containing government and nongovernment unclassified translations relevant to intelligence needs. The CTS functions as a reference check for organizations contemplating a given translation, and it saves the U.S. Government approximately $1 million annually by avoiding duplication. Still further down the road--possibly in FY-82 and FY-83--FBIS plans to automate its Contract Service Order (CSO) system, which monitors the flow of translation from our in-house scanners to JPRS, to our JPRS contract translators, to final publication in the JPRS translations series. STAT STAT STAT

SUBJECT: Seminar on Machine Translation

6. With respect to the basic concept of machine translation per se we admit a substantial degree of skepticism, which our 1976 and 1978 seminars and our subsequent contacts with the MT field have augmented. We know of no breakthrough in the field of artificial intelligence and are confident that had there been any we would have heard about it. Beyond this, however, we believe not only that nothing has happened but that nothing will happen--at least, not in the commonly accepted sense of the purpose of such efforts: a machine system which can effectively replace the human being. On the other hand, such research can clearly have serendipitous benefits in the area of machine-assisted translation, and it is in that area that we see potential benefits in undertaking some form of review.

7. While seminars have their uses in providing forums where disagreements and misconceptions can be resolved and new developments reported, they also have potential for the opposite results: the exchange of unevaluated esoterica, disputed and exaggerated claims, and the temptation to engage in hucksterism. As a major element in the translating community, however, we continue to see a need for continuing contact with the state of the art in both fields--in machine translation, for the byproducts; in machine-assisted translation, for evaluation of the proliferating number of new products (word processors, minicomputers, optical character readers, facsimile transmitting systems, and interfaces of such systems with satellite transmission systems). We suggest that such areas could be more effectively resurveyed by means of a contract with one or two recognized experts who would be given guidelines similar to those given [redacted] in 1976 and 1978 and who would be required to submit reports and evaluations responsive to those guidelines from acknowledged leaders in the business, academic, and government worlds. FBIS would be happy to participate with other interested Agency components and other U.S. agencies in developing guidelines answering our operating needs. We have tried the seminar route; now we would like to see whether the constraints of written reports could channel the discussion into a more usable form. STAT

8. I have asked [redacted] Chief, Production Group, to respond to further questions you may have on any of the points raised above. STAT

[redacted] STAT

Attachments:
A. [redacted] Bid STAT
B. Seminar's Summation

SUBJECT: Seminar on Machine Translation

DDS&T/FBIS/PROD GP/ [redacted] 7 Dec 81 STAT

Distribution:
Original - Addressee, w/attachments
2 - D/FBIS, w/attachments
1 - C/Prod Group, w/attachments
1 - FBIS Registry, w/attachments
1 - Prod Chrono, w/o attachments

ATTACHMENT A

M R M Incorporated
311 Connecticut Avenue
ensington, Maryland 20795

(301) 946-7007

January 21, 1976

[redacted] STAT

Chief, Production Group
Foreign Broadcast Information Service
P. O. Box 2604
Washington, D. C. 20013

Dear [redacted] STAT

In response to your letter of December 12, 1976 and request for proposal at our meeting January 14, 1976, MRM, Inc. wishes to submit the attached proposal to provide services to advise on your selection of speakers and organization of the program and to coordinate and manage all physical arrangements and reservations for an FBIS Seminar on Machine Translation to be held in the Rosslyn area early March, 1976.

Sincerely yours,

[redacted] STAT

President

Attachment

JM/rn

# PROPOSAL

# to

# ADVISE ON PLANNING

# and to

# ADMINISTER AN M-T SEMINAR

MRM, Inc.
January 22, 1976

# Introduction

A two day seminar, sponsored by FBIS is planned to review the applicability of existing machine translation systems to FBIS needs and to establish the state of the art and its potential usefulness to government components. The seminar should include descriptions of currently marketed machine translation systems, survey of the state of research and development programs that bear on present and future machine translation or machine aided translation and the needs of current and potential users.

A conference of this magnitude would normally require six months to plan and execute, in a normal fashion, however, with the liberal use of telephone communication and ready access to the sponsor's task force, we feel that we can provide the necessary services to achieve the objectives. We propose to outline the tasks and cost required to plan and execute the seminar based on an estimated number of speakers and participants and an estimated format for the conduct of the seminar.

It is intended that a two day seminar be conducted in a hotel in the Rosslyn, Virginia area in early March, 1976. The Rosslyn area is desirable because it is convenient and easily accessible to local attendees, and it provides less temptation for diversion from attendance at all conference sessions.

Since the seminar is expected to draw much interest, the physical facilities will be planned for 100 participants and observers and, if necessary, attendance at one or more sessions may be restricted. It may also be desirable for technical reasons to restrict attendance at some sessions to control the efficient development of information for the sponsors.

Speakers recommended for the conference will represent the business community, the academic applied research community, and the government translation community. It will be necessary to provide transportation, housing and per diem expenses for those speakers who do not have resources to cover their own expenses.

Moderators for the seminar will be required to exercise considerable tact and control, especially where claims are challenged and counter claims forcefully made. It will be necessary to carefully select the moderator or moderators for their ability to control the sessions and adequately sum up information of use to the sponsor.

It is recommended that the following guidelines be given to the speakers to insure maximum information gathering:

1) Ask for abstract presentation since time may be too short for copy of complete presentation.
2) Restrict presentation to scientific and technically provable claims.
3) Avoid exaggerated claims for universality, cost, etc.
4) Do not include total machine development as cost if machine can be used for other purposes.
5) Request report of conference from each invited guest.

The planning of the conference and selection of topics and speakers, under the severe time constraint imposed, will require constant attention and frequent consultation with the sponsor. The following tasks and cost data are predicated upon the assumption that the sponsor's task force will be available for consultation

and rapid decision on topics, speakers, and alternates, where necessary.

Costs reflect the services and materials to advise on planning and to administer the seminar, the U. S. travel, accommodations, and consulting fees for a moderator, the U. S. travel and accommodations for 12 invited speakers, and accommodations for 3 guest speakers.

TASKS

Coordinate Seminar

Explore resources of information

Draft agenda

Define probable speakers, moderator, and date

Review and finalize agenda

Search and contact participants

Obtain abstracts

Prepare handouts

Coordinate principal participants

Administer conference

Select conference site

Operate support office

Operate message center

Adminster schedules, rooms, travel, library, attendance

Obtain conference aids - visual, auditory aids, recording, public address, typing reproduction, etc.

Administer registration

Administer travel and reimbursed expenses.

# COST

| | | | | |
|---|---|---|---|---|
| **Travel** | | | | |
| 12 invited speakers | | | | |
| 8 - West Coast | @$350.00 | $2800.00 | | |
| 2 - Midwest | @ 200.00 | 400.00 | | |
| 2 - East Coast | @ 100.00 | 200.00 | | |
| 1 Moderator | | | | |
| 4 trips to D. C. (East Coast) | @ 100.00 | 400.00 | | $3,800.00 |
| **Hotel (12 invited and 3 guest speakers)** | | | | |
| 15 sleeping rooms | 2 days @ $36.00 per day | | $1,080.00 | |
| 1 conference room | 2 days @ $190.00 per day | | 380.00 | |
| 1 conference office | 2 days @ $60.00 per day | | 120.00 | |
| | | | | 1,580.00 |
| **Per diem (15 speakers)** | | | | |
| 15 @ $20.00 per day for 2 days | | | | 600.00 |
| **Conference support equipment (copy, typing, audio/visual, recording)** | | | | 600.00 |
| **Moderator services** | | | | |
| 5 days @$200.00 per day (3 days consulting and 2 days conference) | | | | 1,000.00 |
| **Planning, contacting, coordinative, administrative, and advisory services** | | | | |
| Senior advisor - 1 man month | | $2,500.00 | | |
| Conference assistants - 1 man month | | 850.00 | | |
| Total direct labor | | | $3,350.00 | |
| Overhead @60% | | | 2,010.00 | |
| Fee 12% | | | 643.00 | |
| | | | | 6,003.00 |
| **Total** | | | | $13,583.00 |

ATTACHMENT B

STAT

STAT

# ABSTRACT

FBIS Seminar
(1976)

Very few significant changes in operating machine translation systems have prevailed over the last ten years. All the machine translation programs described produced output suitable for some users when the systems were provided with adequate dictionaries and when context was constrained.

Success of MT is dependent upon further research in linguistics and artificial intelligence. In the mean time applications of advancing technologies such as edit and word processing stations can make considerable impact on all translation operations.

A prima facie case was made for a gradual introduction of languag processing responsive to cumulative technical developments. Such computer-based systems should reflect a reasoned division of tasks which assigns to the computer those storage, sort, and search functions it performs well and reserves for the human the selection, edit, and decision functions appropriate to his skills

STAT

# SUMMARY COMMENTS

## FBIS Seminar
## (1976)

Compiled by [redacted] STAT

Clearly, it is no simple task to evaluate the quality of output achievable through the use of a particular MT system, to determine the amount of post-editing necessary to bring it up to required standards of quality, and to estimate the likely cost of achieving that quality. Each prospective user of an MT system must carefully do this, but from what was presented at this conference, I would not expect any current MT systems to compete economically with human translation except on those few cases where requirements for quality and accuracy are so low that unedited or very lightly edited output suffices.

In addition to postedited MT, this conference also discussed the use of hardware and software aids to human translation. There seemed to be a consensus that well-engineered systems can be produced now, that their use looks promising, and that they probably are limited to increasing the productivity of human translators by a factor of 2-1 or 3-1. Opinion was divided as to whether they might evolve into human-aided MT systems.

[redacted] STAT

The machine translation systems demonstrated were clearly not yet able to completely replace human translation. Whether or not presently available machine translation systems would be useful to an agency's translators in the preparation of human translations would have to be determined by each agency, based on the kind of text, the MT system available, and the type and quality of translation desired as a finished product.

The technology relevant to machine-aided translation is advancing and many costs are coming down. The conclusion is that in order to be prepared for future developments, any agency seriously involved with translation should begin to be involved with this technology, if only on a small scale.

[redacted] STAT

A prima facie case has been made for gradual introduction of language-processing capacity into intelligence facilities.

The main developmental track for a few years ahead is from character processing (editing systems) to word processing (dictionaries).

MT systems, which make no provision for editorial intervention between the earliest and latest stages of processing, are not suitable bases for machine-aided (editorial) systems; and the latter are not necessarily suitable bases for full-scale language-processing systems that may reach installability in as little as ten years if research and development are well-supported.

STAT

We should consider moving MT to microcomputers, to be used along the lines proposed by [redacted]--namely, that the best we can STAT hope to achieve in the immediate future is a kind of "translator's assistant." If existing MT programming ideas could be moved to microcomputers which would perform fast dictionary lookups, rudimentary syntactic analysis and provide convenient text manipulation, then individual research installations could afford to have local copies of a translation rather than waiting on a central installation to perform their translations. In other words, by distributing current MT capabilities--in the form of "intelligent" MT terminals--to the individual researchers who require them for specific classes of translation, the burden on the existing central facilities would be greatly relieved. Microcomputer technology is far enough along to make this entirely reasonable, in the sense that it is nothing more than an engineering issue which could be solved within a year or two.

Specifically, all the systems still rely on basically a word to word translation, with "some" syntax gratuitously tossed in, and dressed up with so-called "micro-glossaries". No system apparently copes with (1) reference, (2) even the most rudimentary model of what it is that's being translated, (3) contextual usage of words (or even automatic microglossary shifting), (4) semantics (e.g., as [redacted] uses it), or (5) "understanding". Basically, STAT the quality of translation is almost directly a function of the quality of the dictionary and how cleverly it has been tuned to a particular application. While this might produce results adequate for a very technical translation to audiences who already know most of what is being said, it falls about ten badly-needed theoretical advances short of so-called "quality" MT, or even low quality

MT in less technical areas, such as literature or political science. In short, there is every reason to be disappointed with the showing.

STAT

With reference to the value of the computer, we should be wary of constructing very elaborate, computer-based systems which do some one or two things very nicely, but which have no generality and make no contribution to the cumulativeness which we must have if we're going to move toward any "utopia" (to use a word employed yesterday) re natural language understanding and applications, such as machine translation. In other words, when we build computer systems, we should think less about ad hoc demonstrations of notions or theories and more about testable, generalizable systems.

I'd urge that a system to be used in machine translation either provide larger screens or keep a kind of running summary which could be used to alert the translator through underlining, a warning message, or whatever, that, for example, a given word or phrase was being used too often.

STAT

One very interesting idea proposed by [redacted] and seconded by several of the participants was that of "human-aided machine translation" whereby one would profit from the human's enormous pattern-recognition and deductive abilities interactively during the translation process. STAT

For the moment, however, text input seems to be a crucial bottle-neck in the use of currently available machine-based translation systems.

In conclusion, while ideal translation by machine of utterances in a source language to the "best fit" in a target language remains a long-range, perhaps unattainable goal, nonetheless enormous practical results appear to be attainable through the development of "human-assisted" programs.

STAT

We feel that the problem of machine translation is only secondarily a matter of computational skill, and remains primarily a linguistic problem.

STAT

The most significant news to emerge from the two days of the FBIS Seminar was that there now are working machine translation systems.

The same brute force approach which enables current systems to work, however, makes it scientifically and linguistically uninteresting. No new linguistically significant generalizations, no new insights in linguistic theory emerge.

STAT

cheerful admission that he is happy when translations cost him as little as $10 per word reflects exactly the right spirit, I believe. Efficiency and comprehensiveness are not significant considerations for the scientific approach to this problem. Research programs such as the ones outlined by and by , aiming STAT incremental approaches to automatic translation, should be critically analyzed to determine the scientific import of each incremental stage before the research is actually undertaken.

STAT

All of the machine translation programs described at the conference are capable of a certain level of proficiency in translation when provided with adequate dictionaries and constrained contexts (through the specification of glossaries). As evidence by the Wright-Pattterson group, this performance is sufficient for use by individuals who are (a) specialists in the technical area of the document and (b) familiar with the odd constructions frequently produced by the machine.

With respect to the more general question of the possible use of computer technology for the sponsor's efforts, the application of Word Processing Stations seems promising. STAT description of what can be done now with advanced graphic editing systems was accurate and encouraging. Furthermore, the use of <u>such computer editors allows a natural evolution</u> to more and more <u>powerful machine aids including automatic translation</u> as they become feasible.

In conclusion, on the short term time scale, the issue which seems worth investigating by the sponsor is whether word processing stations are becoming cost-effective as aids to the in-house translator, especially where the translation is required relatively quick.

STAT

There was consensus that the computer should be introduced into the translation process wherever it is economical for current needs. This includes:

1) use of off-the-shelf items,

2) research and development where off-the-shelf items are not really adequate to the tasks, or

3) establishment of plans for future incorporation of functions which have been developed in the research community but not yet incorporated into off-the-shelf hardware.

It was concluded that during the process of selecting systems or hardware for application to tasks that maximum flexibility and orderly growth be the principal criteria applied to assure economic, long-term usefulness and avoid reliance on costly replacement.

STAT